Efficient Data Dissemination through a Storageless Web Database

Thomas Hammel

prepared for DARPA SensIT Workshop

15 January 2002

Fantastic Data

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

Web Database System

- **Automatically establish redundant data caches throughout the network based on**
  - **data usage patterns, transactions and queries**
  - **optimize cost function based on power consumption, latency, and survivability**
  - **no permanent storage**
- **Disseminate data and maintain redundant caches**
  - **reliable delivery on top of an unreliable channel**
  - **retries mitigated by**
    - **data expiration**
    - **obsolescence detection**
    - **priority**
  - **supports dynamic filter changes**
  - **cooperative repair**

Outline

- **Major focus of last period**
- **What cache does for you**
- **Near term plans**
  - **2 data dissemination techniques**
- **Application Example**
- **Suggested System Block Diagram**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

Major focus of last period

- **Data cache development**
- **Code deliveries**
- **SITEX02**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

Major focus of last period: cache development

- **higher speed, smaller code size**
- **implemented more data types and functions**
- **implemented distributed table creation**
- **implemented “results extraction” filters**
- **interfaces to**
  - **ISI diffusion**
  - **Sensoria radio**
  - **UDP**
  - **simulator**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

Major focus of last period: code deliveries

- **1.1 8 June 2001**
- **1.2 20 August 2001**
- **1.3 9 September 2001**
- **1.4 8 October 2001**

version 1.4 is part of baseline system

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

Major focus of last period: SITEX02

- **supported operational demonstration**
  - **data cache (version 1.4) part of standard system**
  - **data collected through special ISI diffusion processes**
  - **data supplied to UMd gateway for VT GUI and imager triggering**

- **development experiment**
  - **desired real sensor detections from dense deployment of nodes in intersection**
  - **network not ready, so no data collected**
  - **will use data from linear road and simulation**

What cache does for you

- **Data storage**
  - **producer and consumer do not need to directly communicate**
  - **easy to arrange data replay for test and debug**
- **Data dissemination**
  - **all data is available for remote access**
  - **filters automatically determined to satisfy application queries**
- **Primary key for data naming**
  - **automatic consistency enforcement**
  - **data stream merging**
- **Multiple access methods**
  - **real time notification of changes**
  - **search and extract (by query with where clause)**
- **Partial access to structures**
  - **subset of fields**
- **Implemented safely as a server for multiple clients**

Publish/subscribe

- **standard concept in distributed database systems**
  - **10+ years**
  - **off-the-shelf products available since mid 90s**
- **open questions are**
  - **scalability, efficiency, and reliability of dissemination (especially on poor networks)**
  - **filter (subscription) changes**
  - **automatic determination of filters (subscriptions) based on usage**

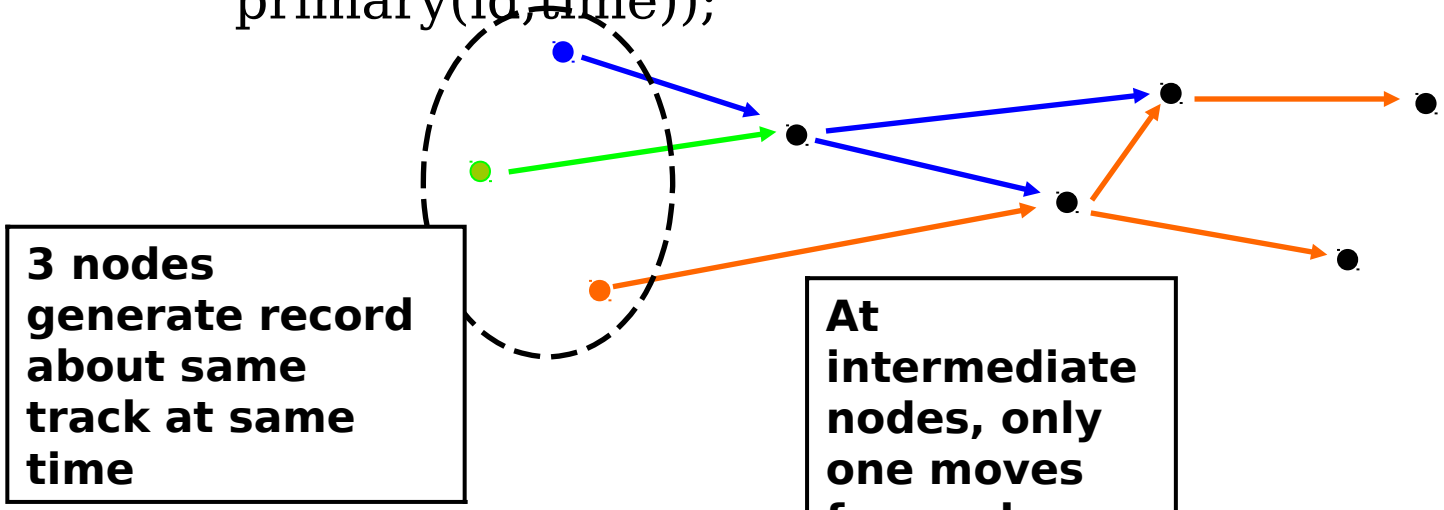
- **how is this implemented in data cache**
  - **subscribe = “watch” query (version 1.x, SITEX02), all queries (version 2.x, coming soon)**
  - **publish = not explicit (version 1.x), may disseminate some statistics in support of 1-time queries (version 2.x)**
  - **evaluated on each record individually**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

Data naming

- **primary key in application's name space**
  - **correct naming allows merging of data streams enroute**

create table track (id c, time u32, … ,
primary(id,time));



- **sequence number in database's name space**
  - **used for bookkeeping**
  - **consistency enforcement**
  - **retransmission and repair**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

Reliable vs. unreliable delivery

- **reliable is too expensive**
  - **data can become obsolete while the system is still trying to deliver it**
- **unreliable is, well, unreliable**
  - **most application programmers don’t want to (can’t) deal with not getting expected data**

- **cache implements guaranteed delivery for data as long as it remains valid**
  - **uses redundant paths through network**
  - **may send multiple times if link reliability is low**
  - **not all data will be delivered, some will become obsolete or expire before delivery**
  - **data will not necessarily be delivered in order**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

Near term plan

- **Support additional users**
- **Implement configuration options**
- **Need to factor 1-time queries and updates into filters**
- **Dissemination filtering techniques**
- **Investigate relationship of data cache to other work**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

Near term plan: Support users

- **Add synchronous operation function call**
  - **not recommended, but it is a lot easier**
- **Figure out what's happening with C++.**
  - **Seems to be 1 operation lag.**
- **Reduce startup bandwidth usage**
  - **In simulation starting 40 nodes simultaneously, usage is about 2KB/s for first 2 minutes. Then drops. Why?**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

Near term plan: Configuration options

- **Data criticality**
  - **cache (1.x) sends all records to all neighbors that are closer to the destination than its own node**
  - **cache (2.x) want to reduce redundancy for less important data items**
- **Latency requirements**
  - **cache (1.x) sends data in order changed**
  - **cache (2.x) is deadline is missed, lower record's place in output queue**
- **Excess data holdback (don't need it more frequently than ...)**
  - **cache (1.x) sends data when communication channel is available**
  - **cache (2.x) send updated record only after certain elapsed time allowing channel to be completely idle**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

Near term plan: 1-time query support

- **Need to factor 1-time queries and updates into filters**
  - **How often are they done?**
  - **How closely do they match the persistent queries?**
  - **How large is the remote load required to satisfy the query?**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

Near term plan: relationship to other work

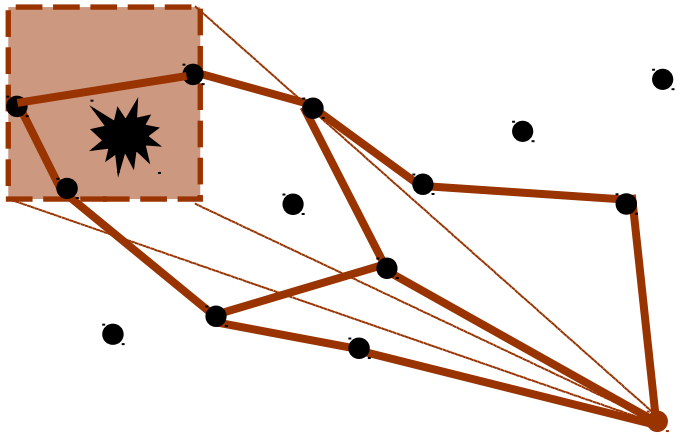
- **Investigate relationship of Fantastic Data caches to ISI routing**
  - **Should we place a filtering module inside the routing layer?**
  - **What are the similarities/differences between our filtering approach and ISI's.**
- **Investigate relationship to Cornell Cougar**
  - **support for in network caching, meta-data, ...**
- **Possibility of direct link to ISI-E mobile GUI**
  - **link through Cornell-Postgres established for baseline demo**
- **Others**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

2 different dissemination problems

Results Formation

- Dense, connected interests
- Data disseminated to neighbors
- Cheap
  - Local broadcast
  - Neighbor's interest can be approximated by own

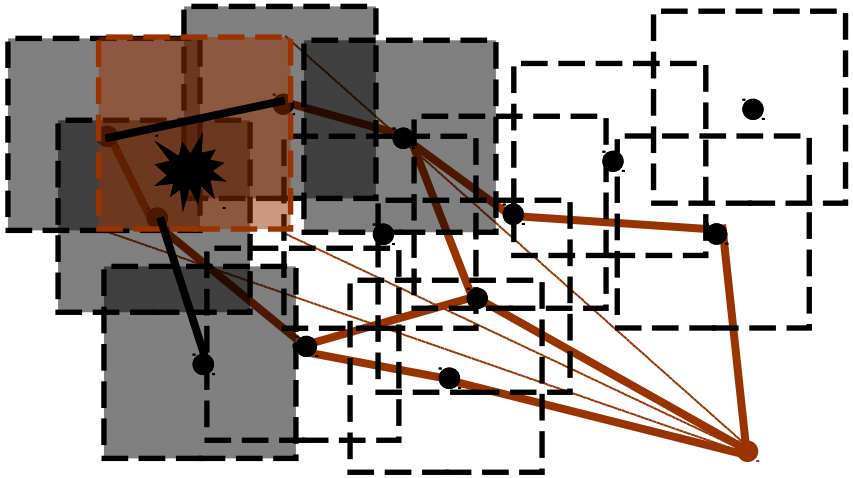


Results Extraction

- Sparse, disjoint interests
- Data moves across network through many uninterested nodes
- Expensive
  - Routing required
  - Requires knowledge of and evaluation of all nodes' interests
- Also satisfies "formation" case at much higher cost

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

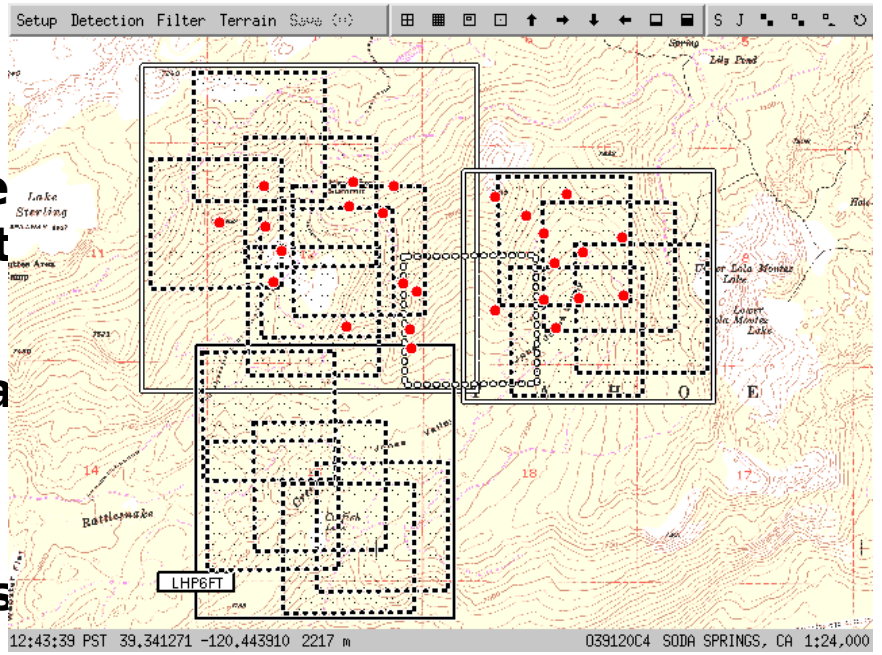
Results Formation and Extraction



- **Implemented both techniques**
- **configured for “extraction” to support UMd gateway and VT GUI**
- **Can we regain efficiency of “formation” technique while still correctly supporting “extraction”?**
  - **“extraction” method needs filter scope reduction to allow network growth**
    - **clustering**
    - **aggregation**
    - **suppression**

Clustering philosophy

- **locally determined**
  - **not globally optimized**
  - **minimize interaction betwe nodes required to setup filt**
- **incremental**
  - **try to disturb existing situa as little as possible**
- **filter tolerance**
  - **a little too big, a little too s that's ok**



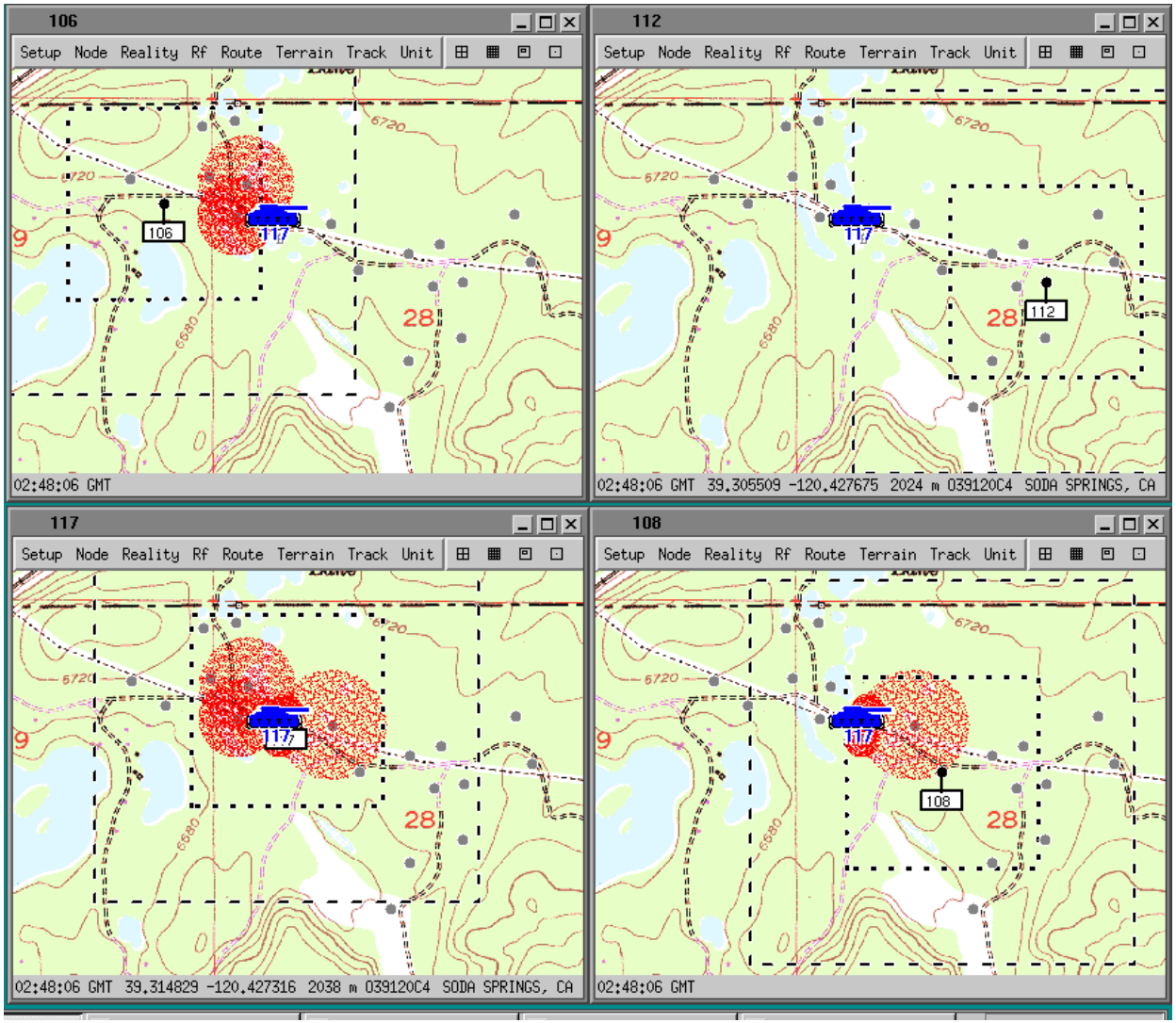
- **maintain cluster quality information**
  - **mean coverage of individual needs (percent, record count, bandwidth)**
  - **excess coverage (percent, record count, bandwidth)**
  - **number of members in group**
  - **mean age of member's input data (seconds)**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • ▯ (415) 643-9554 • ▯ www.fandat.com

Filter scope reduction

- **Suppress filters early in distribution if they are very similar to neighbors**
  - **don’t distribute unless it looks like an “extraction” filter**
  - **treat these few “extraction” filters as special cases**
  - **process using the the normal “formation” technique with a few special cases**
- **Aggregate filters from nodes on the “left” and advertise the composite to the “right”**
  - **distribute different filter to “left” and “right” sides of node**
  - **questions:**
    - **how do we determine “left” and “right”?**
    - **how much impact (overhead) is caused by changing network conditions?**

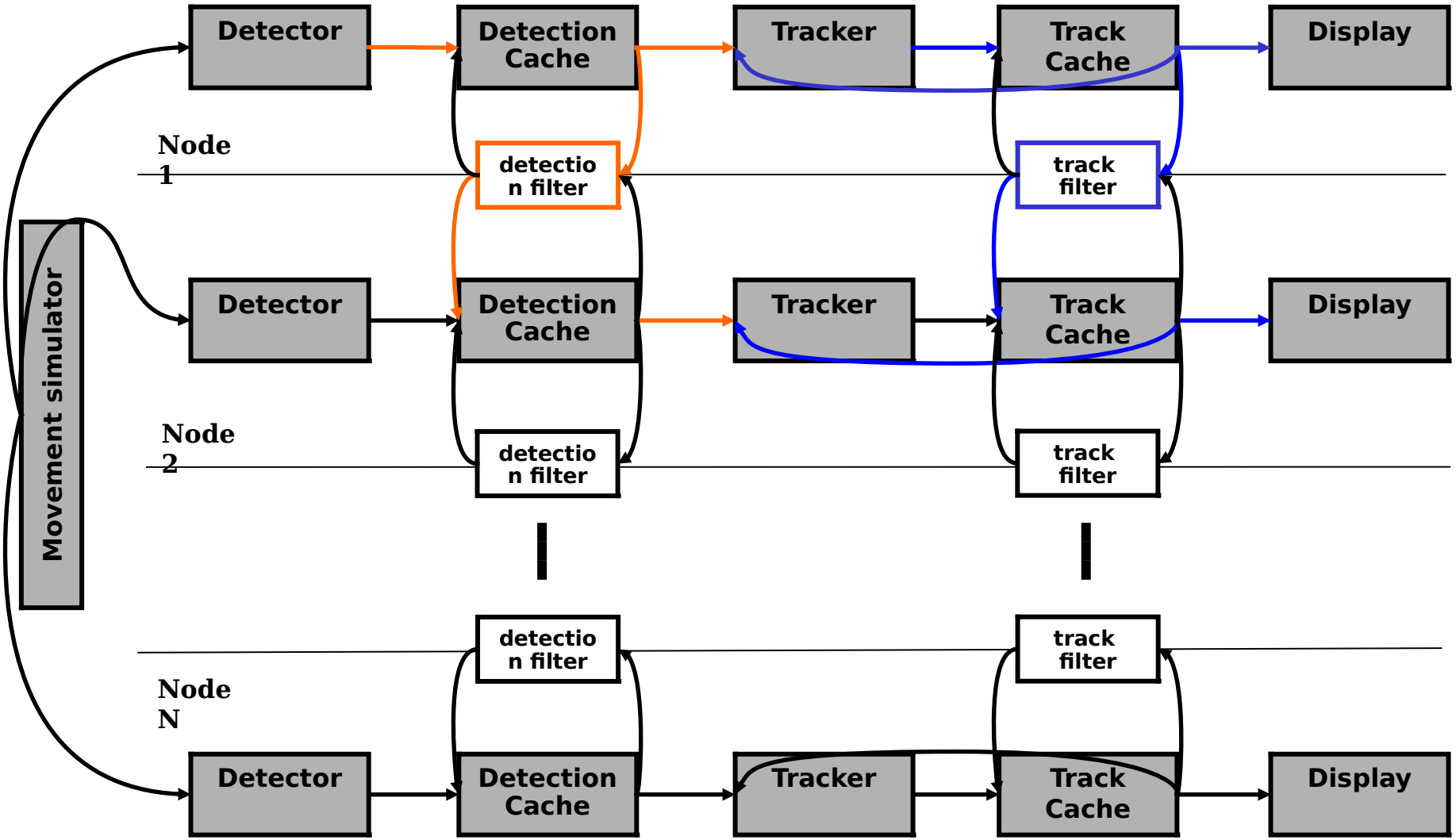
Example from April 2001


106
Setup Node Reality RF Route Terrain Track Unit
02:48:06 GMT
112
Setup Node Reality RF Route Terrain Track Unit
02:48:06 GMT 39.305509 -120.427675 2024 m 039120C4 SODA SPRINGS, CA
117
Setup Node Reality RF Route Terrain Track Unit
02:48:06 GMT 39.314829 -120.427316 2038 m 039120C4 SODA SPRINGS, CA
108
Setup Node Reality RF Route Terrain Track Unit
02:48:06 GMT

Fantastic Data

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • ▯ (415) 643-9554 • ▯ www.fandat.com

Detection and Tracking Example



✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • 🖷 (415) 643-9554 • 🖳 www.fandat.com

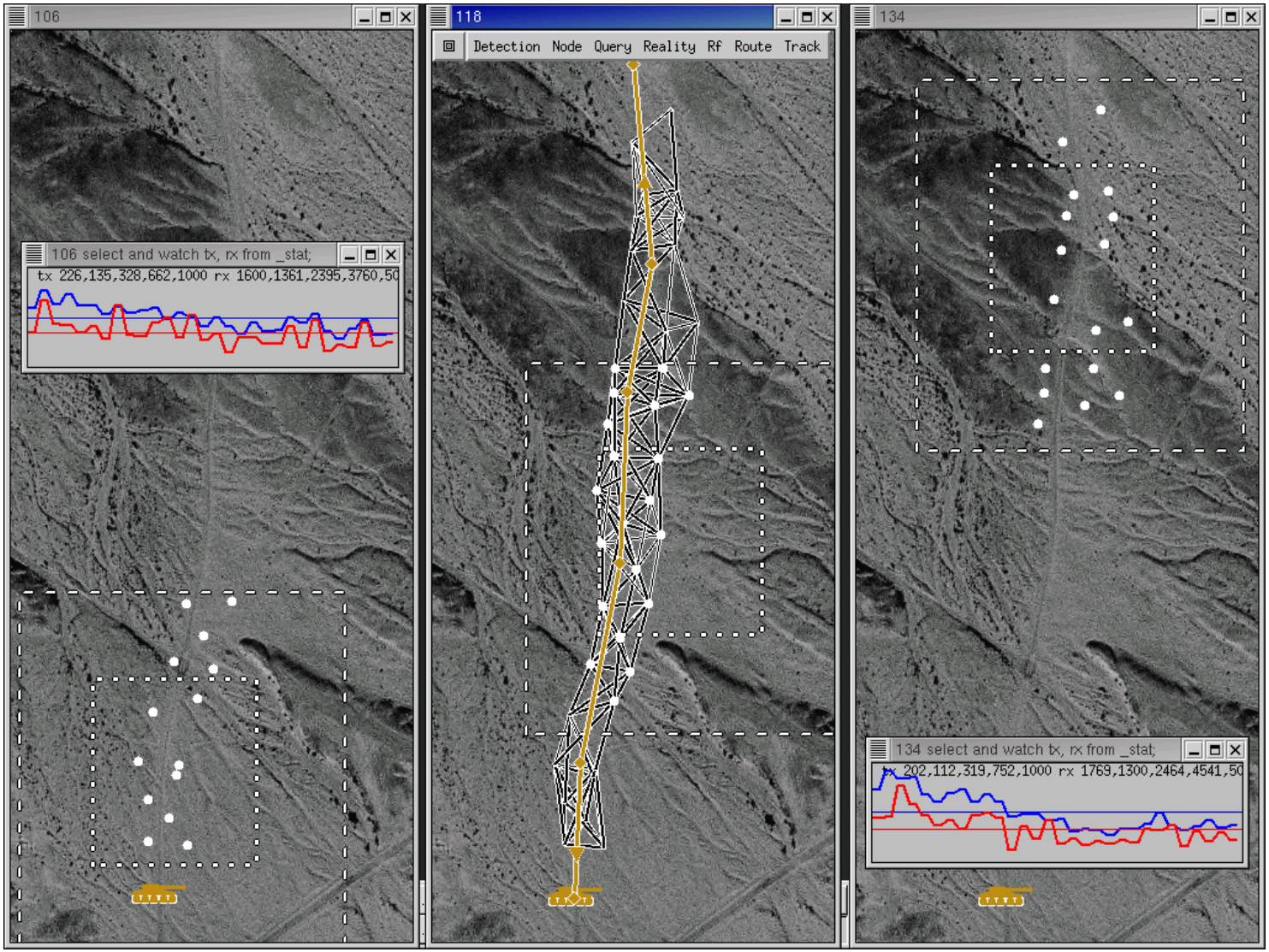
Assumptions

- **No prior knowledge of node locations**
  - **node location is based on gps simulation, averaged over time, and disseminated by the node upon significant change**
- **No prior knowledge of node topology**
  - **neighbors are discovered through broadcast**
  - **link table is computed and disseminated by the nodes**
- **Low power (~10 mW), r^4.3 propagation loss, range is about X m**
- **Poor time synchronization**
  - **node clocks are intentionally off by up to 0.2 seconds**

- **No knowledge of the road**
- **PD is very high, PF very low, detection range about 50 m**
- **Target density is low (>100 m spacing), speed is moderate (<150 Km/h)**
- **Tracking algorithm is a simple data window and least squares fit**

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com

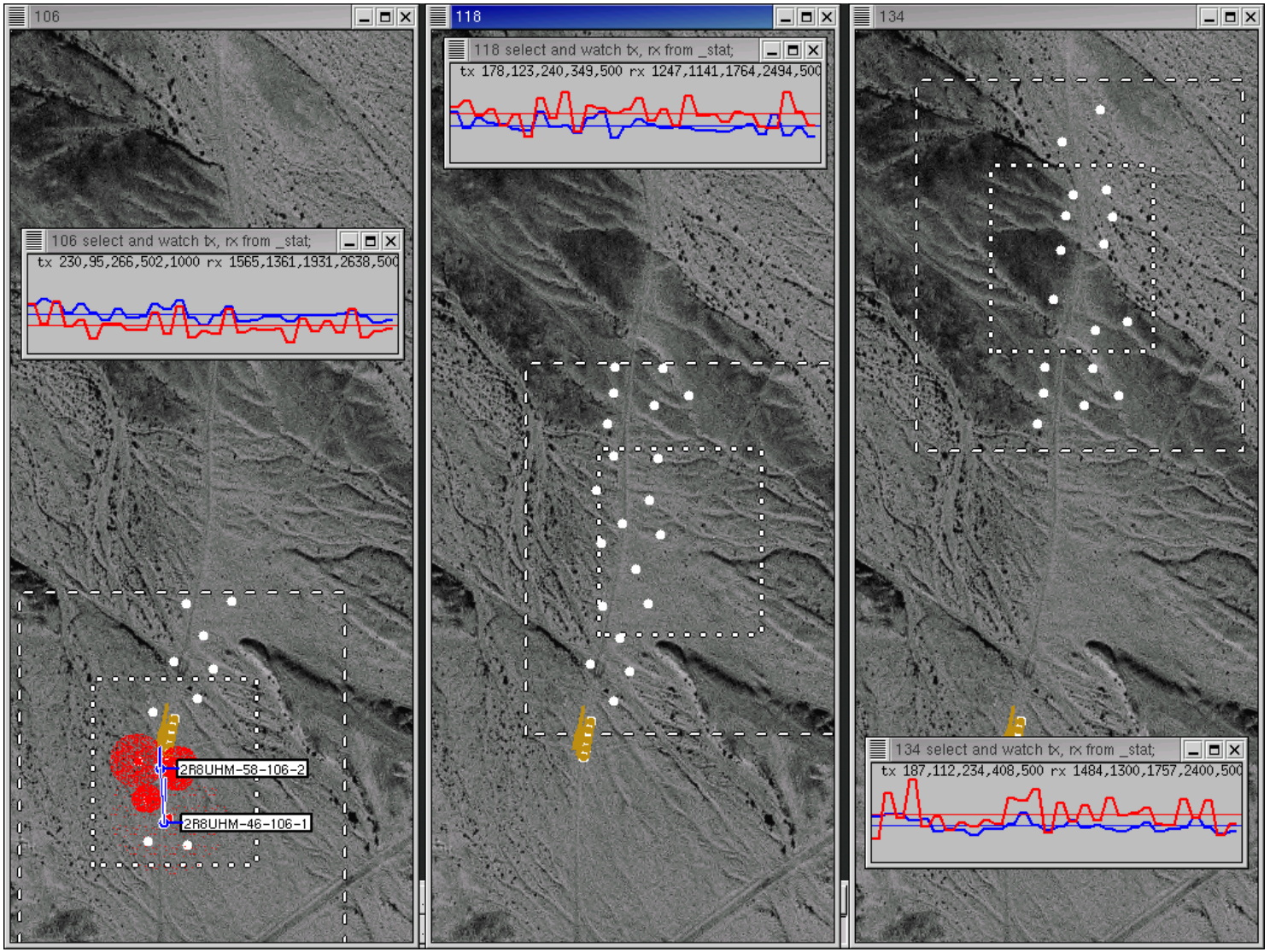
Node Laydown


106
106 select and watch tx, rx from _stat;
tx 226,135,328,662,1000 rx 1600,1361,2395,3760,50
118
Detection Node Query Reality Rf Route Track
134
134 select and watch tx, rx from _stat;
202,112,319,752,1000 rx 1769,1300,2464,4541,50

Fantastic Data

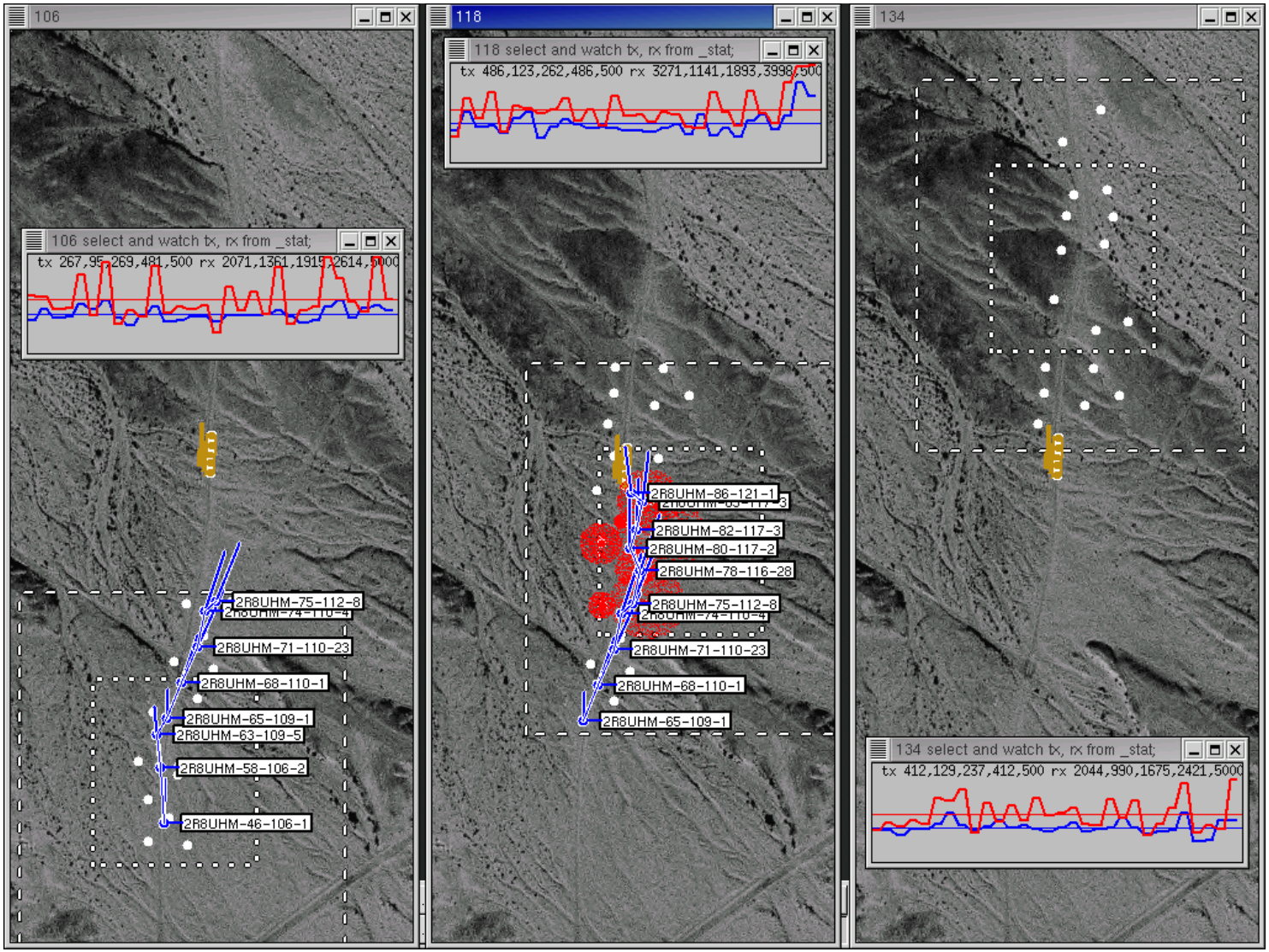
✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • ▯ (415) 643-9554 • ▯ www.fandat.com

Tracking Snapshot 1



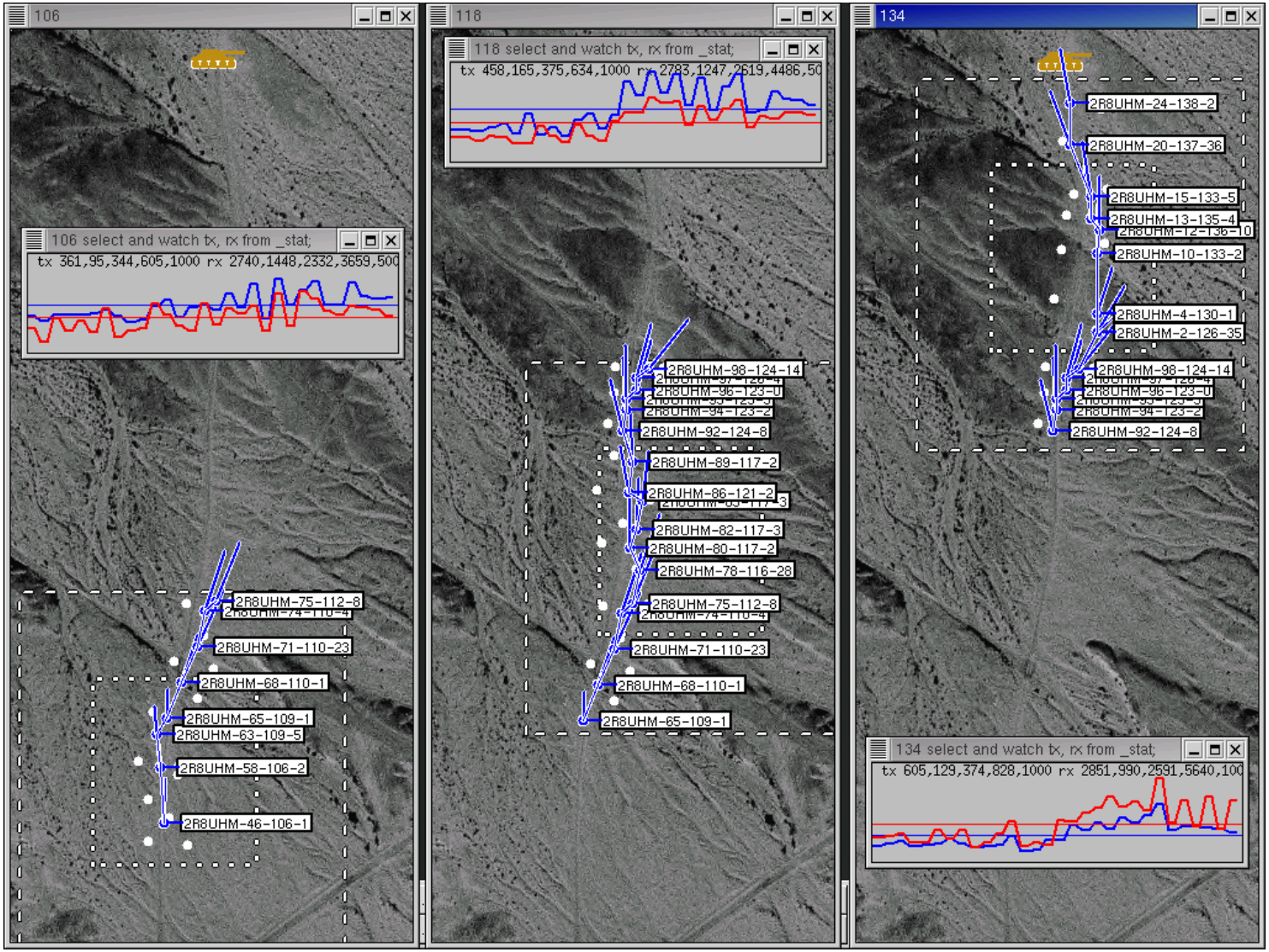
Fantastic Data

Tracking Snapshot 2



Fantastic Data

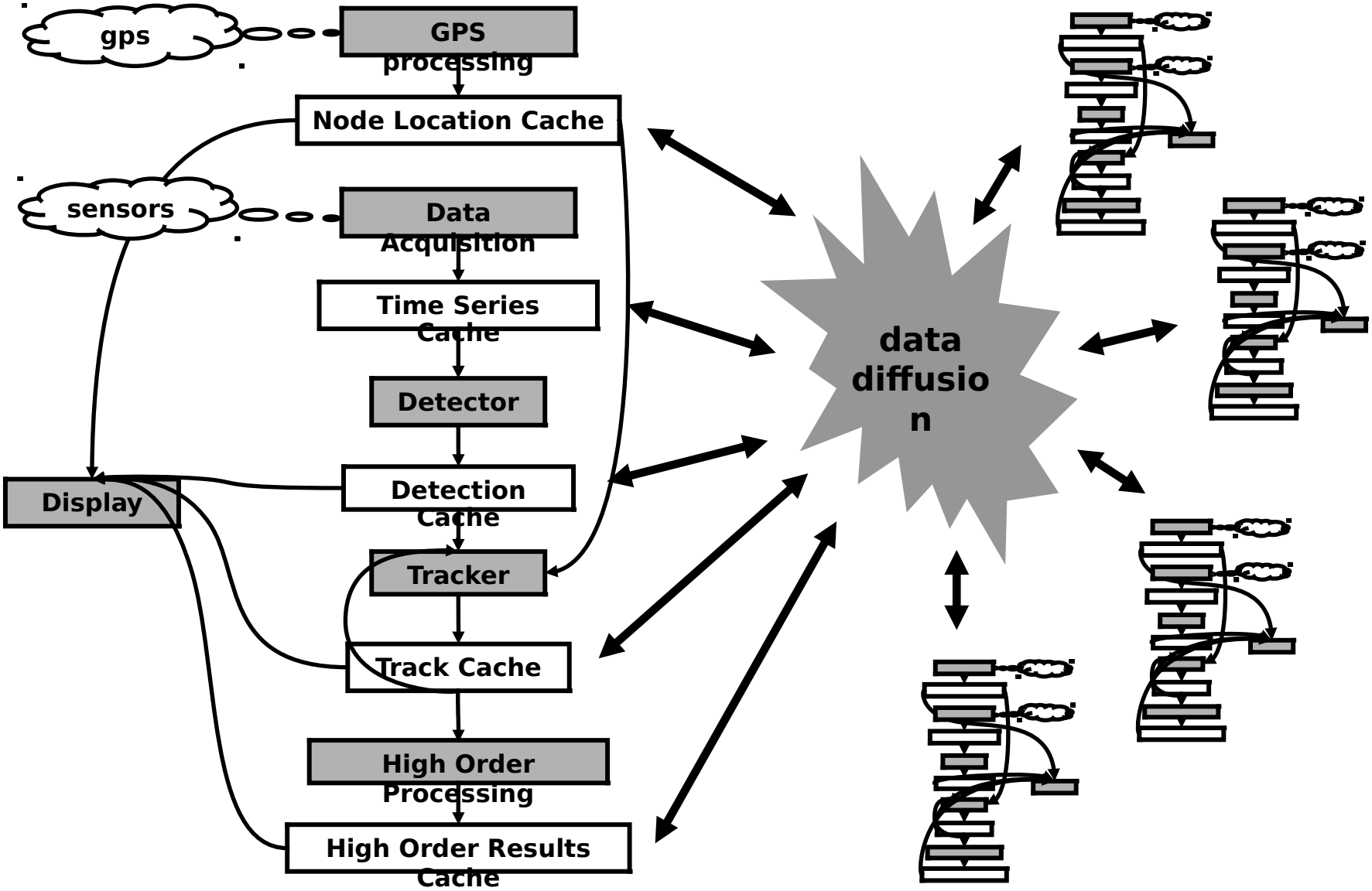
Tracking Snapshot 3



Fantastic Data

✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • ▯ (415) 643-9554 • ▯ www.fandat.com

Suggested System Block Diagram



✉ 207 Prospect Avenue, San Francisco, California 94110, USA • ☎ (415) 643-9555 • (415) 643-9554 • www.fandat.com